

Precision Mars Entry Navigation with Atmospheric Density Adaptation via Neural Networks

Felipe Giraldo-Grueso* and Andrey A. Popov† and Renato Zanetti‡
The University of Texas at Austin, Austin, TX 78712

Spacecraft entering Mars require precise navigation algorithms capable of accurately estimating the vehicle's position and velocity in dynamic and uncertain atmospheric environments. Discrepancies between the true Martian atmospheric density and the onboard density model can significantly impair the performance of spacecraft entry navigation filters. This work introduces a new approach to online filtering for Martian entry using a neural network to estimate atmospheric density and employing a consider analysis to account for the uncertainty in the estimate. The network is trained on an exponential atmospheric density model, and its parameters are dynamically adapted in real time to account for any mismatch between the true and estimated densities. The adaptation of the network is formulated as a maximum likelihood problem by leveraging the measurement innovations of the filter to identify optimal network parameters. Within the context of the maximum likelihood approach, incorporating a neural network enables the use of stochastic optimizers known for their efficiency in the machine learning domain. Performance comparisons are conducted against two online adaptive approaches, covariance matching and state augmentation and correction, in various realistic Martian entry navigation scenarios. The results show superior estimation accuracy compared to other approaches, and precise alignment of the estimated density with a broad selection of realistic Martian atmospheres sampled from perturbed Mars-GRAM data.

I. Introduction

AMONG the most demanding challenges encountered in spacecraft navigation is atmospheric entry on Mars. The entry phase is characterized by intense dynamics, a scarcity of available measurements, and uncertain atmospheric information [1]. Due to the uncertainty in atmospheric conditions, discrepancies between the true atmospheric density and onboard models can make navigation during entry challenging and significantly degrade guidance solutions, affecting landing accuracy [2]. Despite improvements in the knowledge of the Martian atmosphere, its high variability, coupled with a scarcity of comprehensive data, makes it less accurately predictable compared to Earth's atmosphere [1],

*Ph.D. Student, Department of Aerospace Engineering and Engineering Mechanics; fgiraldo@utexas.edu

†Postdoctoral Fellow, Oden Institute for Computational Engineering and Sciences; andrey.a.popov@utexas.edu.

‡Associate Professor, Department of Aerospace Engineering and Engineering Mechanics, AIAA Associate Fellow; renato@utexas.edu

complicating the development of standard Entry, Descent, and Landing (EDL) systems. Standard EDL systems typically use a correction mechanism that adjusts an exponential density profile with a multiplicative factor based on the drag force experienced by the spacecraft [3]. However, constraining density estimation to an exponential profile can pose challenges for Martian entry. For example, the significant atmospheric dust content on Mars is known to increase the temperature of the lower atmosphere and subsequently reduce its density [4]. As a result, future human and robotic Mars exploration missions require improvements by several orders of magnitude over the current state-of-the-art [5].

Previous Mars entry missions have used inertial measurement units (IMU) for navigation instead of onboard models of non-conservative forces. However, IMU-based solutions tend to drift over time, especially without external measurements. If standalone IMU measurements are used for state correction rather than propagation, most estimation states can become unobservable during entry [6]. Therefore, various techniques have been explored to improve navigation solutions during Martian entry. Zanetti and Bishop [7] successfully quantified the accuracy of state estimates through dead-reckoning during standard Martian atmospheric entry only using IMU measurements for navigation. Following this approach, Lou et al. [8] presented a strategy attempting to mitigate the effects of unobservable uncertain parameters through consider analysis during entry. Additionally, Jiang et al. [9] introduced a more accurate multisensor strategy, where instead of using only an IMU for navigation, information from a real-time flush air data system, a single orbiting radio beacon, and an IMU were fused using an unscented Kalman filter, providing better estimation results. Recently, the integration of aerodynamic pressure and aerothermal sensor data for real-time navigation has been proposed [10]. However, most of these solutions face a common challenge: they either rely on an onboard exponential atmospheric density model or depend on external measurements, such as orbiting or surface beacons. The assumption of an exponential density model may pose difficulties when faced with real density profiles, potentially leading to filter divergence. Furthermore, due to the presence of a heat-shield cover on the entry vehicle, onboard optical or radio navigation may not be available during entry [10], making external measurements unavailable.

The descent and landing phases are typically supported by sensors, such as altimeters, velocimeters, or LIDARs, which can aid navigation solutions [11, 12]. However, these sensors are not available during the entry phase, which poses the most significant challenge for navigation. Accurate density estimation becomes crucial in the absence of additional sensors. Therefore, effectively estimating atmospheric density and quantifying its uncertainty can significantly enhance navigation outputs, improving precision landing performance. A range of methodologies have been investigated to address atmospheric uncertainties within the Martian environment with the aim of improving the performance of EDL systems. These approaches can be categorized into two main groups: offline reconstructions, which involve working with real data acquired during the entry phase, and online estimation techniques, which are designed for use during the entry phase. Dutta et al. [13] analyzed various estimation techniques, including the extended Kalman filter (EKF) and unscented Kalman filter (UKF) for offline reconstruction. In their work, the main objective was to reconstruct the

vehicle trajectory, aerodynamic coefficients, and atmospheric profile experienced by the system in an offline context. Subsequently, in their follow-up work [14], they demonstrated how adaptive filtering techniques such as covariance matching [15, 16] could enhance the estimation performance of aerodynamic coefficients and atmospheric properties. These advanced techniques were later employed to successfully reconstruct atmospheric properties using real data acquired by the Mars Science Laboratory (MSL) [17].

To improve filtering precision in online estimation for EDL, various strategies have been explored. A common solution involves carrying a nominal density onboard the filter and augmenting the system state to estimate a correction factor that accounts for the deviation between the nominal density and actual density experienced by the system [2, 10]. In addition, machine learning approaches have been used for density estimation in EDL problems. Wagner et al. [18] used an ensemble of neural networks to dynamically adapt to real-time changes in atmospheric properties during Martian entry. Additionally, Amato and McMahon [19] employed a long short-term memory network to estimate atmospheric density based on inertial measurements and guidance commands. Most recently, Roelke et al. [20] introduced two density profile prediction techniques for aerocapture, using density histories computed from navigation measurements to enhance targeting accuracy.

In this work, we propose a new approach to onboard adaptive filtering for Martian entry by using a neural network, trained pre-flight on a simplified model retrieved from Mars-GRAM trajectories, to estimate atmospheric density. To address the atmospheric mismatch between the true density and the neural network, the network's parameters are adapted in real time using measurement innovations. The measurement innovations are defined as the difference between the true measurements and the predicted measurements of the navigation filter. Furthermore, we account for the uncertainty in the atmospheric density estimate using consider analysis. Consider analysis acknowledges errors in the dynamic and measurement models due to uncertain parameters [21], such as atmospheric density in this case. This approach enables the navigation filter to correctly weight the data while accounting for the correlation between the state and the uncertain parameter [22]. The adaptation of the network is performed using aerodynamic pressure measurements, aerothermal sensor data, and an IMU, mirroring sensor configurations from prior Martian EDL missions [23–25]. Leveraging the capabilities of state-of-the-art stochastic optimizers, nonlinear activation functions, and automatic differentiation, the adaptation step is posed as a maximum likelihood (ML) problem, resulting in consistent filtering performance and accurate estimation of the true density.

The remainder of this paper is organized as follows: First, a background of Martian entry is provided in Section II. Previous approaches to density adaptation are introduced in Section III, and the approach to density adaptation using neural networks is presented in Section IV. The performance of the new approach is compared to previous approaches on a realistic Mars entry navigation scenario in Section V. Section VI discusses some of the limitations of the framework presented and outlines future work opportunities. Finally, Section VII presents the conclusions of the work.

II. Martian Entry

This section provides insight into the dynamics, measurement, and atmospheric density models used in Martian entry. The aim of this analysis is to emphasize the importance of acquiring an accurate atmospheric density estimate during the entry phase.

A. Dynamics

The dynamics of entry vehicles during Martian entry are often represented using a simplified three degrees-of-freedom (3-DOF) model. In this approach and for the entirety of this work, point mass dynamics are assumed, meaning that no consideration is given to attitude dynamics, a common convention observed in prior literature [6, 10, 26–28]. The position of the vehicle is described in terms of planet-centric radius (r), latitude (ϕ), and longitude (θ) which are shown with respect to the Mars-centered Mars-fixed (MCMF) frame [6] in Fig. 1. The MCMF frame is centered and fixed to the planet aligning M_z with the axis of rotation. These coordinates allow the use of spherical equations of motion, even for non-spherical planets, as the radius can vary with latitude and longitude [27]. The dynamic equations used to model the evolution of these variables with respect to time are given by [6, 10, 28],

$$\dot{r} = v \sin(\gamma), \quad (1)$$

$$\dot{\phi} = \frac{\cos(\gamma) \cos(\psi)}{r} v, \quad (2)$$

$$\dot{\theta} = \frac{\cos(\gamma) \sin(\psi)}{r \cos(\phi)} v. \quad (3)$$

The magnitude of the velocity vector (v), flight path angle (γ), and heading azimuth (ψ) are used to describe the velocity of the vehicle as shown with respect to the Geographic frame [6] in Fig. 1. The Geographic frame is centered on the vehicle and G_x , G_Y , G_z are aligned with Mars geographic directions North, East and Nadir, respectively. The flight path angle is defined as negative-down, and the heading azimuth is defined in the horizontal plane with 0° as due North and 90° as due East. The equations describing the time evolution of the velocity-related terms are given by [6, 10, 28]:

$$\dot{v} = -D - g \sin(\gamma) + \omega^2 r \cos(\phi) [\sin(\gamma) \cos(\phi) - \cos(\gamma) \sin(\phi) \cos(\psi)], \quad (4)$$

$$\dot{\gamma} = \frac{1}{v} \left[L \cos(\sigma) - g \cos(\gamma) + \frac{v^2}{r} \cos(\gamma) + 2\omega v \cos(\phi) \sin(\psi) \right] + \frac{1}{v} \left[\omega^2 r \cos(\phi) (\cos(\gamma) \cos(\phi) + \sin(\gamma) \sin(\phi) \cos(\psi)) \right], \quad (5)$$

$$\dot{\psi} = \frac{1}{v} \left[L \frac{\sin(\sigma)}{\cos(\gamma)} + \frac{v^2}{r} \cos(\gamma) \sin(\psi) \tan(\phi) - 2\omega v (\tan(\gamma) \cos(\phi) \cos(\psi) - \sin(\phi)) \right] + \frac{1}{v} \left[\frac{\omega^2 r}{\cos(\gamma)} \sin(\phi) \cos(\phi) \sin(\psi) \right], \quad (6)$$

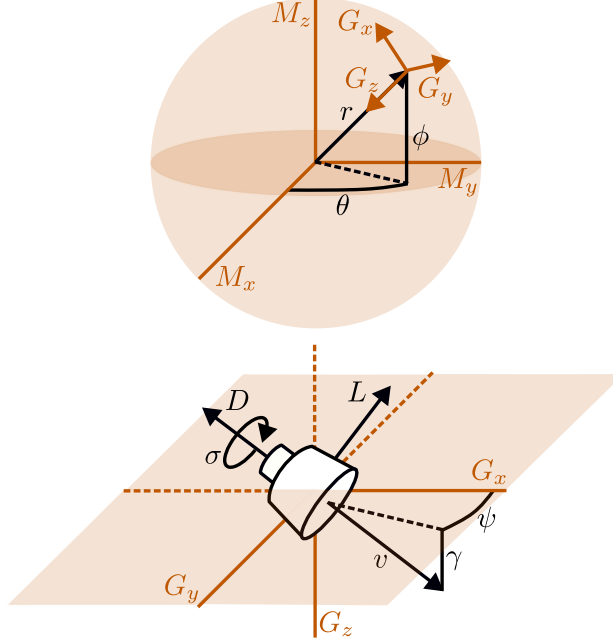


Fig. 1 Position coordinates with respect to the MCMF frame M (top) and velocity coordinates with respect to the Geographic frame G (bottom).

where ω represents the planetary rotation rate, g is the planetary gravity, σ is the bank angle, and the lift (L) and drag (D) accelerations are defined as follows [6]:

$$D = \frac{1}{2} \rho(\cdot) v_{\infty}^2 \frac{C_d S}{m}, \quad (7)$$

$$L = \frac{1}{2} \rho(\cdot) v_{\infty}^2 \frac{C_l S}{m}, \quad (8)$$

with $v_{\infty} = v - W$, where W represents the wind velocity. The parameter $\rho(\cdot)$ is the atmospheric density, where the notation (\cdot) signifies its dynamic nature during the entry phase. Given the dependency of lift and drag accelerations on atmospheric density, all velocity-related terms become directly influenced by changes in atmospheric density. As a result, the temporal evolution of the position implicitly relies on atmospheric density variations.

Given the substantial hypersonic speeds associated with the entry phase, wind velocities are considerably lower compared to the velocity of the vehicle, i.e., $v/W \ll 1$ [19]. In such cases, it is reasonable to make the approximation that $v_{\infty} \approx v$. Moreover, for brief-duration flights such as the entry phase in EDL, the impact of planetary rotation can be deemed negligible [10]. Note that neglecting planetary rotation makes frame M a simple inertial frame. Hence, by omitting terms related to planetary rotation and assuming a first-order gravity model, the equations governing velocity

parameters simplify to [29]:

$$\dot{v} = -\frac{1}{2}\rho(\cdot) v^2 \frac{C_d S}{m} - \frac{\mu}{r^2} \sin(\gamma), \quad (9)$$

$$\dot{\gamma} = \frac{1}{v} \left[\frac{1}{2}\rho(\cdot) v^2 \frac{C_l S}{m} \cos(\sigma) - \frac{\mu}{r^2} \cos(\gamma) + \frac{v^2}{r} \cos(\gamma) \right], \quad (10)$$

$$\dot{\psi} = \frac{1}{v} \left[\frac{1}{2}\rho(\cdot) v^2 \frac{C_l S}{m} \frac{\sin(\sigma)}{\cos(\gamma)} + \frac{v^2}{r} \cos(\gamma) \sin(\psi) \tan(\phi) \right]. \quad (11)$$

By reviewing the dynamics that govern Martian entry, it becomes evident that errors in the estimation of atmospheric density can significantly impact the dynamics, leading to a considerable degradation in the propagation of filtering states. This emphasizes just how crucial it is to attain an accurate estimate of density for navigation solutions.

B. Onboard Sensors

The measurement model of the onboard sensors typically used in EDL missions is closely tied to the atmospheric density estimate as well. In this work, three distinct onboard sensors are under consideration. These sensors encompass an IMU, a cluster of pressure sensors strategically placed on the heat shield of the vehicle, and a suite of thermocouples situated on the forebody of the vehicle. These sensors are carefully selected to replicate the sensor configurations used in past Martian EDL missions [10, 23–25].

1. Inertial measurement unit

The IMU measures the non-gravitational accelerations encountered by the entry vehicle. In this work, these measurements are conducted relative to the body frame aligned with the axes of symmetry in the vehicle [10]. The acceleration measurements are described by:

$$\tilde{\mathbf{a}}^b = T_v^b \mathbf{a}^v + \boldsymbol{\eta}_a, \quad (12)$$

where $\boldsymbol{\eta}_a \in \mathbb{R}^3$ represents measurement noise, T_v^b is the transformation matrix to change coordinates from velocity frame to body frame, and $\mathbf{a}^v \in \mathbb{R}^3$ are the non-gravitational accelerations in the velocity frame defined as [10]:

$$\mathbf{a}^v = \begin{bmatrix} -D & L \sin(\sigma) & L \cos(\sigma) \end{bmatrix}^T. \quad (13)$$

The velocity frame is centered on the vehicle, with its x -axis aligned in the direction of the flight path angle, its z -axis aligned with the local vertical plane orthogonal to the x -axis, and the y -axis completing the right-handed system [10].

2. Cluster of pressure sensors

The pressure sensors are assumed to offer a reasonably accurate evaluation of the stagnation-point pressure. It is important to acknowledge that the precision of these measurements may vary according to the flight regime. For instance, pressure transducers optimized for hypersonic flight may not deliver satisfactory results during supersonic flight. Nevertheless, this work assumes the availability of measurements throughout the entire entry stage. The dynamic pressure measurement, assuming an aggregate measurement from all relevant sensor arrays, is expressed as follows [10]:

$$\tilde{q} = \frac{1}{2} \rho(\cdot) v^2 + \eta_q, \quad (14)$$

where $\eta_q \in \mathbb{R}$ represents measurement noise.

3. Suite of thermocouples

The suite of thermocouples is assumed to provide real-time measurements of the convective heating rate. The measured convective heating rate, assuming again an aggregate measurement from all relevant sensor arrays, is modeled by the Sutton-Graves relation [30, 31],

$$\tilde{Q}_s = k \left(\frac{\rho(\cdot)}{R_n} \right)^{\frac{1}{2}} v^3 + \eta_{\dot{Q}_s}, \quad (15)$$

where $k = 1.9027 \times 10^{-4} \text{ kg}^{1/2} \text{ m}^{-1}$, R_n is the vehicle nose radius and $\eta_{\dot{Q}_s} \in \mathbb{R}$ is measurement noise.

These equations further emphasize the critical importance of having an accurate atmospheric density estimate, as they all rely on this fundamental parameter. This connection illustrates how easily a navigation system can diverge when confronted with inaccurate estimates or inaccurate density models stored in the onboard computer of the vehicle, emphasizing the role of obtaining a highly accurate density estimate for mission success and safety.

C. Density Models

To model Martian atmospheric density, various approaches are available, ranging from more intricate models for increased accuracy to simpler equations for faster estimates, although with reduced precision.

1. Exponential models

The most straightforward equation for modeling atmospheric density is based on an exponential law, assuming isothermal equilibrium [32]. This model is represented by the equation:

$$\hat{\rho}_{\text{exp}} = \rho_0 \exp\left(-\frac{(r - r_0)}{h_s}\right), \quad (16)$$

where r_0 represents the reference radial position, ρ_0 is the density at the reference radial position, and h_s stands for the atmospheric scale height. It is worth noting that while the atmosphere is not isothermal at all altitudes, the parameters in (16) can be adjusted piece-wise to better align with the real atmospheric density. However, as previously mentioned, constraining Martian density to an exponential profile can be challenging due to the variability in dust content, seasonal changes, and wind patterns [4]. In response to potential errors in the exponential model, the COSPAR model was developed using data collected from Viking 1, Viking 2, and Mariner missions [1]. Although lacking consideration for factors like longitude, latitude, seasonal variations, and dust content, the COSPAR model offers more precision than the basic exponential model. The simplified COSPAR model is expressed as a modified exponential equation,

$$\hat{\rho}_{\text{COSPAR}} = \rho_0 \exp \left\{ -\beta_\rho (r - r_0) + \gamma_\rho \cos [\omega_\rho (r - r_0)] + \delta_\rho \sin [\omega_\rho (r - r_0)] \right\}, \quad (17)$$

where β_ρ , γ_ρ , ω_ρ and δ_ρ are nominal coefficients adjustable to better fit real atmospheric data. Both the exponential and COSPAR models rely on ad hoc parametrizations for estimating atmospheric density.

2. Mars-GRAM

The Mars Global Reference Atmospheric Model (GRAM) 2010 represents a significant improvement in atmospheric modeling compared to previous models. Mars-GRAM is an engineering-level atmospheric model constructed using input data tables derived from NASA Ames Mars General Circulation Model (MGCM) and the University of Michigan Mars Thermospheric General Circulation Model (MTGCM) outputs. This model has undergone validation against Radio Science and thermal emission spectrometer (TES) data, encompassing both nadir and limb observations. Additionally, it can model potential perturbations in the atmosphere, a feature frequently employed in Monte Carlo simulations for high-fidelity engineering [33].

While Mars-GRAM excels in providing realistic estimates for potential atmospheric density profiles encountered during entry missions, it is usually avoided for online estimation due to its substantial computational overhead. Instead, online estimation approaches commonly use simplified models paired with adaptive techniques, validated through Monte Carlo simulations using truth trajectories propagated using atmospheric density values obtained from Mars-GRAM [2].

III. Approaches to Density Adaptation

As models cannot provide a perfect estimation of the atmospheric density, various adaptive techniques can be used to address potential errors or unforeseen behavior that the vehicle may encounter during the entry phase. The first approach involves covariance matching, wherein adjustments are made to the process and measurement noise covariance matrices to align with the state and measurement innovations in the navigation filter. The second technique entails the estimation of correction parameters, used to rectify a nominal density profile established during the offline phases of the

navigation filter design. Lastly, multiple model adaptive estimation can also be used for this purpose, where a bank of estimators operates in parallel, yielding a more robust estimate by considering different density profiles. This section delves into the presentation and discussion of each of these approaches.

A. Covariance Matching

Covariance matching techniques have found utility in atmospheric reconstruction due to the inherent uncertainty in atmospheric parameters, where the process and measurement noise covariances matrices are often unknown. When directly estimating the density, that is, the density is part of the estimation state, covariance matching techniques have been shown to be advantageous [14]. It is worth noting that incorporating the density as part of the state introduces constraints, notably on the dynamics, as the equation for the rate of change is derived from the hydrostatic equation and the perfect gas law [13].

Mehra [15] introduced a covariance matching algorithm yielding a non-unique solution for the estimated covariances. Shortly after, Myers and Tapley [16] presented an algorithm for providing a unique estimate of noise covariance matrices using batches of state and measurement innovations. To estimate the covariance matrices, the use of the following equations was proposed [16]. Considering a linear state estimation problem, that is,

$$\mathbf{x}_k = \Phi_{k,k-1}\mathbf{x}_{k-1} + \mathbf{q}_{k-1}, \quad (18)$$

$$\mathbf{y}_k = H_k\mathbf{x}_k + \boldsymbol{\eta}_k, \quad (19)$$

where, \mathbf{x} represents the state, k is the time index, $\Phi_{k,k-1}$ is the state transition matrix that maps from time step $k - 1$ to time step k , H_k is the state observation matrix, \mathbf{q}_{k-1} and $\boldsymbol{\eta}_k$ represent process and measurement noise realizations respectively, the estimated process noise covariance matrix is given by [16]:

$$\hat{Q} = \frac{1}{N-1} \sum_{i=1}^N \left[\left(\mathbf{v}_i^{(x)} - \bar{\mathbf{v}}_i^{(x)} \right) \left(\mathbf{v}_i^{(x)} - \bar{\mathbf{v}}_i^{(x)} \right)^T - \left(\frac{N-1}{N} \left[\Phi_{i,i-1} P_{i-1}^+ \Phi_{i,i-1}^T - P_i^+ \right] \right) \right], \quad (20)$$

where N is the batch size, and $\mathbf{v}^{(x)}$ are the state innovations, such that:

$$\mathbf{v}_i^{(x)} = \hat{\mathbf{x}}_i^+ - \Phi_{i,i-1} \hat{\mathbf{x}}_{i-1}^+. \quad (21)$$

The (+) represents updated (corrected) estimates and P_i^+ is the posterior state error covariance. Similarly, the estimated measurement noise covariance is given by [16]:

$$\hat{R} = \frac{1}{N-1} \sum_{i=1}^N \left[\left(\mathbf{v}_i^{(y)} - \bar{\mathbf{v}}_i^{(y)} \right) \left(\mathbf{v}_i^{(y)} - \bar{\mathbf{v}}_i^{(y)} \right)^T - \left(\frac{N-1}{N} \left[H_i P_i^- H_i^T \right] \right) \right], \quad (22)$$

where $\mathbf{v}^{(y)}$ are the measurement innovations:

$$\mathbf{v}_i^{(y)} = \mathbf{y}_i - H_i \hat{\mathbf{x}}_i^- . \quad (23)$$

In this case, the $(-)$ represents propagated (predicted) estimates and P_i^- is the propagated state error covariance. Despite the simplicity of covariance matching for adapting to uncertainty in dynamics and measurement models, it is worth noting that the lack of convergence proof [15] can raise concerns, especially if wanted to be used for online estimation in missions requiring high precision, such as Martian entry.

B. State Augmentation and Correction

Another approach to atmospheric density adaptation includes the correction to a nominal profile by augmenting the estimation state [2, 10]. In this solution, a nominal density profile (ρ_{nominal}) is determined offline and carried by the onboard navigation system. The nominal profile is corrected in an online fashion to account for potential errors. To correct the nominal profile, different approaches can be used. For example, the estimated state can be augmented with a correction factor for the density [2],

$$\mathbf{x}_a = [\mathbf{x} \quad \mathcal{K}]^T , \quad (24)$$

where:

$$\mathcal{K} = \frac{\rho}{\rho_{\text{nominal}}} . \quad (25)$$

With this approach, the correction factor can be estimated within the navigation filter and the estimated density can be calculated as:

$$\hat{\rho} = \hat{\mathcal{K}} \rho_{\text{nominal}} . \quad (26)$$

Estimating \mathcal{K} behaves favorably from a numerical perspective, because the estimator primarily seeks to refine an existing approximate density profile rather than directly estimating the density [2]. The estimated value tends to hover near one, regardless of altitude, when the nominal density profile closely approximates the actual density experienced by the vehicle. Previous guidance approaches have estimated this correction factor by using the sensed deceleration measurements of the vehicle and applying a low-pass ad hoc filter [20]. When a Kalman filter is used for navigation, augmenting the filtering state with this correction factor effectively takes care of applying a low-pass filter.

A different way of correcting a nominal profile involves directly modeling the atmosphere as in (16) and estimating the density at the reference radial position (ρ_0) [10, 29] and the atmospheric scale height (h_s), thus,

$$\mathbf{x}_a = [\mathbf{x} \quad \rho_0 \quad h_s] . \quad (27)$$

However, constraining the nominal density model to an exponential form can be ill-conditioned because the exponential model tends to amplify errors at higher altitudes [14]. It is important to note that augmentation approaches perform the adaptation to the atmospheric density as a linear function of the measurement, since the EKF and the UKF are typically used in these scenarios.

C. Others

Lastly, multiple model adaptive estimation [1, 34] and the use of neural ensembles [18] have been proposed for the adaptation of the atmospheric density in Martian entry; these solutions are highly dependent on the number of filters used (in the case of multiple model filtering) or in the number of density models stored in the neural ensemble.

IV. Density Adaptation with Neural Networks

This work introduces a new approach that uses neural networks within a maximum likelihood framework to adapt to fluctuations in atmospheric density. Previous approaches that have employed maximum likelihood adaptive filtering for tasks like flight-path reconstruction [35], typically rely on analytical expressions and Newton-Raphson methods to solve the optimization problem. In contrast, we propose the use of the maximum likelihood approach within a machine learning framework, enabling the use of well-established stochastic optimization methods that have undergone extensive research and refinement [36]. This new adaptive filtering technique follows two distinct stages: offline training of a neural network, and its online adaptation for accurate atmospheric density estimation.

A. Offline Training

First, a neural network is trained to predict atmospheric density as a function of the planet-centric radius, as represented by the equation,

$$\hat{\rho}_{NN} = NN(r, \xi), \quad (28)$$

where the weights and biases of the trained neural network are denoted by ξ . To train the neural network, synthetic data is generated by simulating Martian entry dynamics using a low-fidelity model of atmospheric density. To create the low-fidelity model, a least-squares exponential fit is obtained from multiple realizations of Mars-GRAM profiles. This fitting process aims to match the data to an exponential model represented as (16). Using this exponential model, it becomes straightforward to generate synthetic data by simulating various entry trajectories and recording planet-centric radius and atmospheric density at a predefined sampling rate.

Algorithm 1 provides an overview of the data generation and offline training of the neural network. For this algorithm N trajectories are simulated using the exponential density model and starting from different initial conditions. The planet-centric radius and atmospheric density values are saved at a fixed rate for each trajectory to generate the input and output training data. Section IV.A.1 provides an in-depth description of the offline training implementation.

Algorithm 1 Offline data generation and training

- 1: Calculate exponential fit (16) $\hat{\rho}_{\text{exp}}$ from Mars-GRAM profiles
 - 2: Initialize an input vector as empty $\mathcal{R} = []$
 - 3: Initialize an output vector as empty $\mathcal{P}_{\text{exp}} = []$
 - 4: **for** $i = 1 : N$ **do**
 - 5: Set initial conditions: \mathbf{x}_0
 - 6: Propagate initial conditions until a final time t_f using $\hat{\rho}_{\text{exp}}$
 - 7: Save r_k and $\hat{\rho}_{\text{exp},k}$ from the trajectory at a fixed rate
 - 8: Append the sequence of r_k to the input vector \mathcal{R}
 - 9: Append the sequence of $\hat{\rho}_{\text{exp},k}$ to the output vector \mathcal{P}_{exp}
 - 10: Preprocess input vector $\mathbf{i} \leftarrow \mathcal{R}$
 - 11: Preprocess output vector $\mathbf{o} \leftarrow \mathcal{P}_{\text{exp}}$
 - 12: Train neural network from the preprocessed input and output (\mathbf{i}, \mathbf{o}) : $\hat{\rho}_{NN} = \mathcal{NN}(r, \xi)$
-

1. Practical implementation

This section describes the implementation of the offline training of the neural network. It is important to highlight that various implementations can be employed for the neural network, as this network is only a starting point for the online estimation portion. To generate the training data, as previously described, a least squares exponential fit was applied to multiple Mars-GRAM profiles (ρ_{GRAM}). Figure 2 shows the Mars-GRAM profiles used for the implementation of the neural network. To generate these profiles, the density and wind random perturbations were set to 1, representing a standard deviation in the range of 2% to 45% of the unperturbed mean.

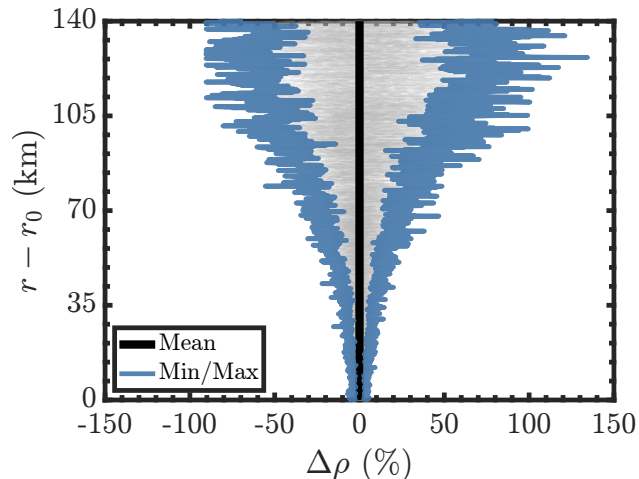


Fig. 2 Altitude as a function of the density variation for the Mars-GRAM profiles generated. The gray lines show the different atmospheric profiles.

Using this exponential fit, 1000 different entry trajectories were simulated based on the dynamic equations outlined in Section II.A. Each trajectory was initialized with distinct initial conditions sampled from the entry conditions described in Table 1, and data on atmospheric density and planet-centric radius were recorded at 0.5-second intervals during the simulations. The trajectories were propagated for 250 seconds reaching an altitude of approximately 11 kilometers above the Martian surface, a typical point marking the end of the entry phase [10]. Figure 3 shows different propagated

trajectories that use different initial conditions coupled with the exponential fit for the density calculation.

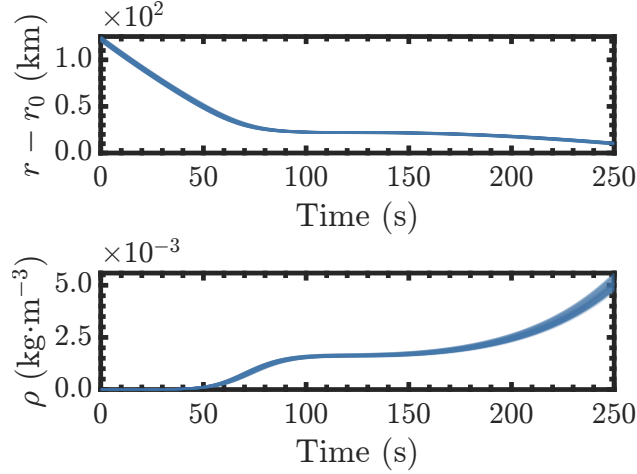


Fig. 3 Altitude (top) and density as obtained from the least squares exponential fit (bottom) as a function of time.

To train the neural network, a method similar to that introduced by Amato and McMahon [19] was adopted. To facilitate the training process, the output data, which represents atmospheric density, was transformed using the equation,

$$\varrho_k = \sqrt{\mathcal{B} - \log_{10} \rho_k}, \quad (29)$$

where \mathcal{B} is a shifting factor ensuring that the argument of the square root is always positive. For this work, $\mathcal{B} = 0$ since the atmospheric density is consistently expressed in SI base units, ensuring that the argument of the square root is always non-negative. To assess the possibility of overfitting, a train/validation split on the data was conducted. Both the training input and output data were normalized, following the transformations,

$$i_k = \frac{r_k - \bar{r}}{\sigma_r}, \quad (30)$$

$$o_k = \frac{\varrho_k - \bar{\varrho}}{\sigma_\varrho}. \quad (31)$$

where i_k is the input to the network and o_k represents the output. Here, \bar{r} represents the mean of all training input data, σ_r is the associated standard deviation, $\bar{\varrho}$ is the mean of all training output data, and σ_ϱ is the associated standard deviation. With the normalized training data, a simple feed-forward network with one hidden layer containing 100 neurons, and using tanh as the activation function, was trained. The architecture of this neural network was chosen to accelerate the adaptation process, aiming for a smaller number of hidden layers and neurons, as the weights and biases of the network will be adapted online and recursively.

The network was trained for 1000 epochs using Adam as the optimizer with a varying learning rate ranging from

$l_{r_{\min}} = 10^{-6}$ to $l_{r_{\max}} = 10^{-2}$. Figure 4 illustrates the procedure for estimating atmospheric density using the trained neural network. The planet-centric radius (r_k) is first normalized, and this normalized input (i_k) is fed into the neural network. Subsequently, the output (o_k) is unnormalized to derive the transformed density (ϱ_k), and the density (ρ_k) is computed using the inverse mapping. The network operations are given by:

$$\mathbf{a}_k = \tanh \left\{ W^{(i)} i_k + \mathbf{b}_k^{(i)} \right\}, \quad (32)$$

$$\mathbf{o}_k = W^{(o)} \mathbf{a}_k + \mathbf{b}_k^{(o)}, \quad (33)$$

where $\mathbf{a}_k \in \mathbb{R}^{100}$ are the neurons in the hidden layer, $W^{(i)} \in \mathbb{R}^{100 \times 1}$ is the vector of weights for the first layer, $i_k \in \mathbb{R}$ represents the input and $\mathbf{b}_k^{(i)} \in \mathbb{R}^{100}$ are the biases for the first layer. Similarly, $o_k \in \mathbb{R}$ is the output of the network, and $W^{(o)} \in \mathbb{R}^{1 \times 100}$ and $\mathbf{b}_k^{(o)} \in \mathbb{R}$ are the weights and biases for the output layer, respectively. Note that the time complexity of a forward pass through the network is $O(1 \cdot h_0 + h_0 \cdot 1) = O(2 \cdot h_0)$ with h_0 the number of neurons (in this case 100).

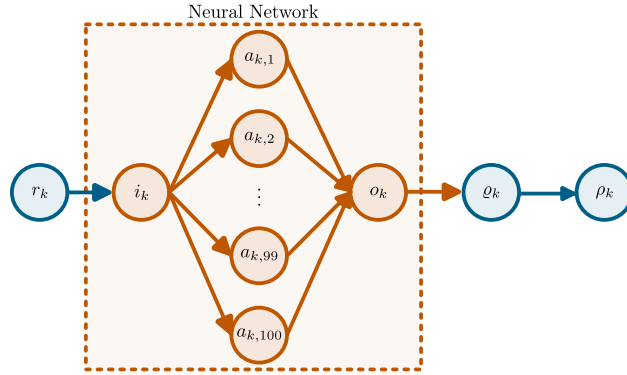


Fig. 4 Network architecture used to estimate the atmospheric density from the planet-centric radius.

Figure 5 displays the percentage error between the estimated density from the trained neural network and the actual density in the validation set. As it can be seen, the network accurately predicts the density, as most errors are less than 1%, suggesting a good fit to the exponential model. It is important to note that the accuracy of this network is not crucial for the proposed approach. The network is expected to perform poorly on real atmospheric density profiles, which is why the weights and biases will be updated online to adapt to a higher fidelity density model. However, achieving a good fit to an exponential model is desirable.

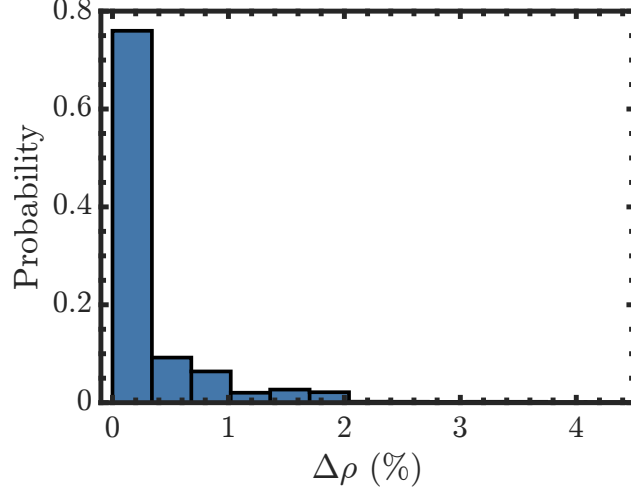


Fig. 5 Probability histogram of the relative error between the predicted density from the trained neural network and the density in the validation set.

B. Online Estimation

Once the neural network has been trained, an online adaptation scheme can be applied. For this work, the state to be estimated is defined as:

$$\mathbf{x} = [r \quad \phi \quad \theta \quad v \quad \phi \quad \psi \quad B \quad L/D]^T, \quad (34)$$

where the inverse of the ballistic coefficient (B) is modeled as a random walk, employing a similar approach for modeling the lift-to-drag ratio (L/D) [10]. The available measurement vector is expressed as:

$$\mathbf{y} = [\tilde{\mathbf{a}}^b \quad \tilde{q} \quad \tilde{Q}_s]. \quad (35)$$

To simplify notation, the dynamics and measurement models are defined as:

$$\dot{\mathbf{x}} = f(\mathbf{x}, \rho), \quad (36)$$

$$\mathbf{y}_k = h(\mathbf{x}_k, \rho_k) + \boldsymbol{\eta}_k, \quad (37)$$

where f are the dynamics presented in section II.A, h refers to the measurement model in section II.B, ρ is the atmospheric density and $\boldsymbol{\eta}_k$ represents measurement noise. Process noise is modeled in the solution of the dynamics by,

$$\mathbf{x}_k = F(\mathbf{x}_{k-1}, \rho_{k-1}) + \mathbf{q}_{k-1} \quad (38)$$

where F is the flow of (36) and \mathbf{q}_{k-1} is the process noise. For this scenario, the selected online estimator is the unscented Schmidt-Kalman filter (USKF) [22, 37]. The USKF is a linear update filter known for its robustness in

handling nonlinear dynamics and measurements. This choice eliminates the need for partial derivatives both during time propagation and measurement updates, which can otherwise lead to computational slowdown as the neural network is being used to approximate the density, thus requiring the need of more complicated partial derivatives. Additionally, this filter incorporates a consider analysis to account for the uncertainty in the atmospheric density estimates, as this quantity is not part of the filter's state space.

1. Time propagation

Each iteration starts from the previous time posterior estimate of the state $\hat{\mathbf{x}}_{a,k-1}^+$ and covariance $P_{a,k-1}^+$, where both the estimate and covariance are augmented to include the consider parameter,

$$\hat{\mathbf{x}}_{a,k-1}^+ = \begin{bmatrix} \hat{\mathbf{x}}_{k-1}^{+\text{T}} & \hat{\mathbf{c}}_{k-1}^+ \end{bmatrix}^{\text{T}}, \quad (39)$$

$$P_{a,k-1}^+ = \begin{bmatrix} P_{k-1}^+ & C_{k-1}^+ \\ (C_{k-1}^+)^{\text{T}} & P_{c,k-1}^+ \end{bmatrix}, \quad (40)$$

where the initial value $P_{c,0}^+$ is user-defined (note that $\hat{\mathbf{c}}_0^+ = 1$ and $C_0^+ = 0_{n_x \times 1}$). With these augmented state and covariance, a set of posterior estimates is built by calculating the associated sigma points [38],

$$\hat{\boldsymbol{\chi}}_{k-1}^{+(0)} = \hat{\mathbf{x}}_{a,k-1}^+, \quad (41)$$

$$\hat{\boldsymbol{\chi}}_{k-1}^{+(i)} = \hat{\mathbf{x}}_{a,k-1}^+ + \left(\sqrt{L + \lambda P_{a,k-1}^+} \right)^{(i)}, \quad i = 1, \dots, L, \quad (42)$$

$$\hat{\boldsymbol{\chi}}_{k-1}^{+(i)} = \hat{\mathbf{x}}_{a,k-1}^+ - \left(\sqrt{L + \lambda P_{a,k-1}^+} \right)^{(i)}, \quad i = L + 1, \dots, 2L, \quad (43)$$

resulting in a set of $2L + 1$ samples. The superscript (i) denotes the i -th column of the matrix in parenthesis. The scaling parameter λ is defined as $\lambda = \alpha^2 (L + \kappa) - L$, where α and κ are secondary scaling parameters and $L = n_x + 1$. The parameter α controls the spread of points around $\hat{\mathbf{x}}_{a,k-1}^+$ and is typically set to a small positive value ($10^{-4} \leq \alpha \leq 1$), and κ is usually chosen as $3 - L$ [38].

After calculating the sigma points, these can be propagated to the next time step by integrating the dynamics and using the trained neural network to estimate the density. With $F_a(\mathbf{x}, \rho)$ representing both the flow of $f(\mathbf{x}, \rho)$ and the flow of the consider parameter modeled as an exponentially correlated random variable (ECRV) [21] centered at one, the states are propagated by (the operator $\mathcal{NN}(\cdot)$ handles the normalization and transformation steps):

$$\hat{\boldsymbol{\chi}}_k^{-(i)} = F_a \left(\hat{\boldsymbol{\chi}}_{k-1}^{+(i)}, \mathcal{P}_{k-1}^{(i)} \right), \quad (44)$$

$$\mathcal{P}_{k-1}^{(i)} = \hat{\boldsymbol{\chi}}_{k-1}^{+(i)+} (n_x + 1) \cdot \mathcal{NN} \left(\hat{\boldsymbol{\chi}}_{k-1}^{+(i)}(1), \boldsymbol{\xi}_{k-1} \right), \quad (45)$$

where $\hat{\mathbf{X}}_{k-1}^{(i)+}(n_x + 1)$ is used as Matlab notation and represents the consider parameter at each sigma point, $\hat{\mathbf{X}}_{k-1}^{+(i)}(1)$ is the estimated planet-centric radius at each sigma point and ξ_{k-1} are the current weights and biases of the neural network. Note that the consider parameter is dispersing the solution of the neural network to account for the uncertainty in the atmospheric density estimate. With the propagated sigma points, a prior estimate and covariance can be calculated by a weighted sum, such that:

$$\hat{\mathbf{x}}_{a,k}^- = \sum_{i=0}^{2L} w_m^{(i)} \hat{\mathbf{X}}_k^{-(i)}, \quad (46)$$

$$P_{a,k}^- = \sum_{i=0}^{2L} w_c^{(i)} \left[\hat{\mathbf{X}}_k^{-(i)} - \hat{\mathbf{x}}_k^- \right] \left[\hat{\mathbf{X}}_k^{-(i)} - \hat{\mathbf{x}}_k^- \right]^T + Q_{a,k-1}, \quad (47)$$

where the weights are given by [39]:

$$w_m^{(0)} = \frac{\lambda}{L + \lambda}, \quad (48)$$

$$w_c^{(0)} = \frac{\lambda}{L + \lambda} + 1 - \alpha^2 + \beta, \quad (49)$$

$$w_m^{(i)} = w_c^{(i)} = \frac{1}{2(L + \lambda)}, \quad i = 1, \dots, 2L, \quad (50)$$

and the augmented process noise covariance is defined as:

$$Q_{a,k-1} = \begin{bmatrix} Q_{k-1} & 0_{n_x \times 1} \\ 0_{1 \times n_x} & Q_{c,k-1} \end{bmatrix}, \quad (51)$$

with $Q_{c,k-1} = (1 - \exp\{-2\Delta t_k/\tau\}) P_{ss}$, where Δt_k is the propagation interval, τ is the time constant of the ECRV and P_{ss} is the steady-state covariance of this variable. The constant β is used to incorporate prior knowledge of the probability distribution. Under the assumption of a scalar and Gaussian prior distribution, $\beta = 2$ is usually regarded as optimal since it can cancel second-order error terms [40]. From the augmented state and covariance, the estimation state and its covariance can be retrieved by selecting the first n_x elements of the resulting vector and the n_x by n_x upper block of the resulting matrix.

2. Maximum likelihood optimization

Once a measurement (\mathbf{y}_k) is obtained, a loss function is constructed using the measurement log likelihood,

$$\mathcal{L}(\xi_{k-1}) = [\mathbf{y}_k - h(\hat{\mathbf{x}}_k^-, \hat{\rho}_{k-1})]^\top R^{-1} [\mathbf{y}_k - h(\hat{\mathbf{x}}_k^-, \hat{\rho}_{k-1})], \quad (52)$$

$$\hat{\rho}_{k-1} = \mathcal{NN}(\hat{\mathbf{x}}_k^-(1), \xi_{k-1}). \quad (53)$$

This loss function highlights the fact that the adaptation is performed in a nonlinear fashion with respect to the measurement, as opposed to the previous augmentation approaches described in section III. Thus, a minimization problem can be constructed, where,

$$\xi_k = \arg \min_{\xi_{k-1}} \mathcal{L}(\xi_{k-1}). \quad (54)$$

When dealing with non-convex problems, fine-tuning the learning rate in traditional optimization algorithms (such as Newton-Raphson) can present challenges and may lead to divergence issues. Instead, we propose the use of automatic differentiation and stochastic optimizers to address the minimization problem in (54). Given that this problem pertains to the optimization of the weights and biases of a neural network, back-propagation (also known as adjoint-mode differentiation) can be employed to compute gradients up to floating-point precision. Furthermore, the adaptive and stochastic optimization capabilities offered by the Adam optimizer [36] can be leveraged to minimize the loss function. A notable advantage of using the Adam optimizer is its ability to incorporate memory of previous gradients from earlier time steps, thereby assisting in the fine-tuning of the adaptive learning rate within the optimizer. Starting from an initial gradient $\nabla_{\xi_{k-1}} \mathcal{L}$, an initial moving average of the gradient $\bar{\nabla}_{k-1}$, and an initial moving average of the squared gradient $(\bar{\nabla}_{k-1})^2$, where the notation $(\cdot)^2$ indicates element-wise square, Adam iteratively updates the network parameters as follows. First the moving average of the gradient and the gradient squared are updated,

$$\bar{\nabla}_{k-1}^* = \beta_1 \bar{\nabla}_{k-1} + (1 - \beta_1) \nabla_{\xi_{k-1}} \mathcal{L}, \quad (55)$$

$$(\bar{\nabla}_{k-1}^*)^2 = \beta_2 (\bar{\nabla}_{k-1})^2 + (1 - \beta_2) (\nabla_{\xi_{k-1}} \mathcal{L})^2, \quad (56)$$

where β_1 and β_2 are exponential decay rates of the moving averages being estimated. Once these two estimates are obtained, the network parameters are updated using the following equation:

$$\xi_{k-1}^* = \xi_{k-1} - \alpha_{\mathcal{L}} \left(\bar{\nabla}_{k-1}^* / \left[\sqrt{(\bar{\nabla}_{k-1}^*)^2 + \varepsilon} \right] \right). \quad (57)$$

Here, $\alpha_{\mathcal{L}}$ represents the step size, and ε is a small constant used to prevent numerical overflow. This process repeats until a convergence criterion is met. When a convergence criterion is satisfied, the updated network parameters and

moving averages are saved for the measurement update and the next iteration. Thus, upon reaching convergence,

$$\bar{\nabla}_k = \bar{\nabla}_{k-1}^*, \quad (58)$$

$$(\bar{\nabla}_k)^2 = (\bar{\nabla}_{k-1}^*)^2, \quad (59)$$

$$\xi_k = \xi_{k-1}^*. \quad (60)$$

For this work, the moving average of the gradient was initialized to zero, and the square of the gradient was initialized to the current gradient squared. Additionally, the moving averages were not corrected for the initialization bias. These design strategies were informed by the performance of the algorithm when deployed online in different filtering tests. In practice, the maximization step is only performed if the evaluated loss exceeds a predetermined threshold (\mathcal{T}). This helps alleviate the computational burden of the online filter and leads to less noisy outcomes. Algorithm 2 provides a comprehensive overview of the maximum likelihood optimization step. In this algorithm, a maximum patience value (p_{\max}) is defined to prevent the degradation of the neural network in the event of encountering unfavorable gradients.

Algorithm 2 Maximum likelihood optimization (MLO)

Input: $\hat{x}_{a,k}^-, P_{a,k}^-, \xi_{k-1}, \bar{\nabla}_{k-1}, (\bar{\nabla}_{k-1})^2, \alpha_{\mathcal{L}}, \beta_1, \beta_2, p_{\max}, \mathcal{T}$

Output: $\xi_k, \bar{\nabla}_k, (\bar{\nabla}_k)^2$

- 1: Calculate initial loss (52): $\mathcal{L}(\xi_{k-1})$
 - 2: Initialize patience: $p = 0$
 - 3: **while** $\mathcal{L}(\xi_{k-1}) > \mathcal{T}$ **do**
 - 4: Apply back-propagation to obtain the gradients: $\nabla_{\xi_{k-1}} \mathcal{L}$
 - 5: Use Adam to obtain new parameters, average gradient and average gradient squared:
 - 6: $\left[\xi_{k-1}^*, \bar{\nabla}_{k-1}^*, (\bar{\nabla}_{k-1}^*)^2 \right] = \text{Adam} \left[\xi_{k-1}, \nabla_{\xi_{k-1}} \mathcal{L}, \bar{\nabla}_{k-1}, (\bar{\nabla}_{k-1})^2, \alpha_{\mathcal{L}}, \beta_1, \beta_2 \right]$
 - 7: Evaluate new loss (52): $\mathcal{L}(\xi_{k-1}^*)$
 - 8: **if** $\mathcal{L}(\xi_{k-1}^*) < \mathcal{L}(\xi_{k-1})$ **then**
 - 9: Update adapted parameters: $\xi_{k-1} \leftarrow \xi_{k-1}^*$
 - 10: Update Adam parameters: $\left(\bar{\nabla}_{k-1}, (\bar{\nabla}_{k-1})^2 \right) \leftarrow \left(\bar{\nabla}_{k-1}^*, (\bar{\nabla}_{k-1}^*)^2 \right)$
 - 11: Update loss: $\mathcal{L}(\xi_{k-1}) \leftarrow \mathcal{L}(\xi_{k-1}^*)$
 - 12: **else**
 - 13: Update patience: $p = p + 1$
 - 14: **if** $p = p_{\max}$ **then**
 - 15: **break**
 - 16: Save adapted parameters: $\xi_k \leftarrow \xi_{k-1}$
 - 17: Save Adam parameters: $\left(\bar{\nabla}_k, (\bar{\nabla}_k)^2 \right) \leftarrow \left(\bar{\nabla}_{k-1}, (\bar{\nabla}_{k-1})^2 \right)$
-

As previously mentioned, the time complexity of a forward pass through the network is $O(2 \cdot h_0)$. Assuming that the back-propagation shares the same order of magnitude in time complexity as the forward pass, the total time complexity for one optimization step is $O(2 \cdot h_0)$. The time complexity scales linearly with the number of optimization steps (n_o), resulting in $O(2 \cdot h_0 \cdot n_o)$. Therefore, careful consideration of the number of optimization steps is important to ensure affordable online implementation. Nonetheless, these time complexity considerations lie outside the scope of the current

research, as the primary goal of this work is to demonstrate the capabilities of using neural networks in an adaptive manner.

3. Measurement update

Once the adapted weights and biases of the network have been obtained, the measurement update can be performed. To do this, sigma points are calculated around the prior estimate by using the propagated state and covariance and using the same approach as in equations (41)-(43). With the set of prior sigma points, a set of expected measurements is calculated by using the measurement model and the adapted weights and biases of the network,

$$\hat{\mathbf{y}}_k^{(i)} = h\left(\hat{\mathbf{x}}_k^{-(i)}, \mathcal{P}_k^{(i)}\right), \quad (61)$$

$$\mathcal{P}_k^{(i)} = \hat{\mathbf{x}}_k^{-(i)}(n_x + 1) \cdot \mathcal{NN}\left(\hat{\mathbf{x}}_k^{-(i)}(1), \xi_k\right). \quad (62)$$

Notice again that the atmospheric density estimate obtained from the network is slightly perturbed by the consider parameter for each sigma point in order to account for uncertainty. With the set of expected measurements, the predicted mean, its covariance and cross-covariance with the state can be calculated to perform a Kalman update. Using a similar weighted sum as in the time propagation,

$$\hat{\mathbf{y}}_k = \sum_{i=0}^{2L} w_m^{(i)} \hat{\mathbf{y}}_k^{(i)}, \quad (63)$$

$$P_{yy,k} = \sum_{i=0}^{2L} w_c^{(i)} \left[\hat{\mathbf{y}}_k^{(i)} - \hat{\mathbf{y}}_k \right] \left[\hat{\mathbf{y}}_k^{(i)} - \hat{\mathbf{y}}_k \right]^T + R_k, \quad (64)$$

$$P_{xay,k} = \sum_{i=0}^{2L} w_c^{(i)} \left[\hat{\mathbf{x}}_k^{-(i)} - \hat{\mathbf{x}}_k^- \right] \left[\hat{\mathbf{y}}_k^{(i)} - \hat{\mathbf{y}}_k \right]^T, \quad (65)$$

where R_k is the measurement noise covariance, and the weights are calculated using the same equations as in the time propagation. It is important to note that for the measurement update, the constants α , β and κ do not necessarily have to be the same as in the time propagation step. With the obtained matrices, a Kalman gain is calculated,

$$K_k = P_{xay,k} P_{yy,k}^{-1} = \begin{bmatrix} K_{x,k} \\ K_{c,k} \end{bmatrix}. \quad (66)$$

Finally, a Kalman update can be performed to obtain the posterior estimate and covariance, such that:

$$\hat{\mathbf{x}}_{a,k}^+ = \hat{\mathbf{x}}_{a,k}^- + \begin{bmatrix} K_{x,k} \\ 0 \end{bmatrix} (\mathbf{y}_k - \hat{\mathbf{y}}_k), \quad (67)$$

$$P_{a,k}^+ = P_{a,k}^- - \begin{bmatrix} K_{x,k} P_{yy,k} K_{x,k}^T & K_{x,k} P_{yy,k} K_{c,k}^T \\ K_{c,k} P_{yy,k} K_{x,k}^T & 0 \end{bmatrix}. \quad (68)$$

Just as with the time propagation, from the augmented state and covariance, the estimation state and its covariance can be retrieved by selecting the first n_x elements of the resulting vector and the n_x by n_x upper block of the resulting matrix. It is important to note that the consider parameter and its associated covariance are not updated in this filter.

Algorithm 3 shows the full recursion algorithm for the online estimation. Initially, the neural network is trained offline to obtain the initial weights and biases (ξ_0) and the initial Adam parameters ($\bar{\nabla}_0, (\bar{\nabla}_0)^2$). In addition, the initial state and covariance ($\hat{\mathbf{x}}_{a,0}^+, P_{a,0}^+$) are set as the starting point ($k-1$). Then, the state and covariance are propagated ($\hat{\mathbf{x}}_{a,k}^-, P_{a,k}^-$) using the current network. Once the measurement is obtained (\mathbf{y}_k), the loss function (\mathcal{L}) is minimized using Adam and the network (ξ_k) and optimizer parameters ($\bar{\nabla}_k, (\bar{\nabla}_k)^2$) are adapted. Finally, the state and covariance are updated ($\hat{\mathbf{x}}_{a,k}^+, P_{a,k}^+$) using the adapted network. A summary of the complete online estimation scheme is also presented in Fig. 6 as a flowchart.

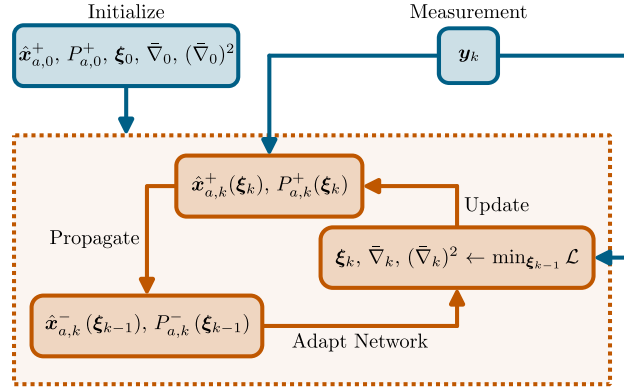


Fig. 6 Flowchart of the online estimation scheme.

V. Results and Discussion

To assess the density adaptation scheme, a comprehensive Monte Carlo testing approach was employed. This involved conducting 1000 Monte Carlo runs to thoroughly evaluate the robustness and performance of the algorithms presented. Each true trajectory was simulated using an independent and distinct Mars-GRAM testing trajectory, different from the set previously mentioned, which was used for obtaining the least squares exponential fit and training the network.

Algorithm 3 Recursion algorithm

- 1: Initialize estimation parameters: $\hat{\mathbf{x}}_{a,0}^+, P_{a,0}^+$
 - 2: Initialize neural network parameters: $\xi_0, \bar{\mathbf{v}}_0, (\bar{\mathbf{v}}_0)^2$
 - 3: Declare optimization variables: $\alpha_{\mathcal{L}}, \beta_1, \beta_2, p_{\max}, \mathcal{T}$
 - 4: Declare unscented transform constants: α, β, κ
 - 5: **while** \mathbf{y}_k is available **do**
 - 6: \triangleright *Time propagation* \triangleleft
 - 7: Calculate sigma points (41)-(43) $\hat{\mathcal{X}}_{k-1}^+$ from $\hat{\mathbf{x}}_{a,k-1}^+, P_{a,k-1}^+$
 - 8: Propagate sigma points to current time step: $\hat{\mathcal{X}}_k^- = F_a \left(\hat{\mathcal{X}}_{k-1}^+, \mathcal{P}_{k-1} \right)$
 - 9: Calculate sigma points weights (48)-(50): $w_m^{(i)}, w_c^{(i)}$
 - 10: Reduce sigma points:
 - 11: $\hat{\mathbf{x}}_{a,k}^- = \sum_{i=0}^{2L} w_m^{(i)} \hat{\mathcal{X}}_k^{-(i)}$
 - 12: $P_{a,k}^- = \sum_{i=0}^{2L} w_c^{(i)} \left[\hat{\mathcal{X}}_k^{-(i)} - \hat{\mathbf{x}}_k^- \right] \left[\hat{\mathcal{X}}_k^{-(i)} - \hat{\mathbf{x}}_k^- \right]^T + Q_{a,k-1}$
 - 13: \triangleright *Maximum likelihood optimization (Algorithm 2)* \triangleleft
 - 14: $\left[\xi_k, \bar{\mathbf{v}}_k, (\bar{\mathbf{v}}_k)^2 \right] = \text{MLO} \left[\hat{\mathbf{x}}_{a,k}^-, P_{a,k}^-, \xi_{k-1}, \bar{\mathbf{v}}_{k-1}, (\bar{\mathbf{v}}_{k-1})^2, \alpha_{\mathcal{L}}, \beta_1, \beta_2, p_{\max}, \mathcal{T} \right]$
 - 15: \triangleright *Measurement update* \triangleleft
 - 16: Calculate sigma points (41)-(43) $\hat{\mathcal{X}}_k^-$ from $\hat{\mathbf{x}}_{a,k}^-, P_{a,k}^-$
 - 17: Calculate expected measurements: $\hat{\mathbf{y}}_k = h \left(\hat{\mathcal{X}}_k^-, \mathcal{P}_k \right)$
 - 18: Calculate sigma points weights (48)-(50): $w_m^{(i)}, w_c^{(i)}$
 - 19: Reduce sigma points:
 - 20: $\hat{\mathbf{y}}_k = \sum_{i=0}^{2L} w_m^{(i)} \hat{\mathcal{Y}}_k^{(i)}$
 - 21: $P_{yy,k} = \sum_{i=0}^{2L} w_c^{(i)} \left[\hat{\mathcal{Y}}_k^{(i)} - \hat{\mathbf{y}}_k \right] \left[\hat{\mathcal{Y}}_k^{(i)} - \hat{\mathbf{y}}_k \right]^T + R_k$
 - 22: $P_{xay,k} = \sum_{i=0}^{2L} w_c^{(i)} \left[\hat{\mathcal{X}}_k^{-(i)} - \hat{\mathbf{x}}_k^- \right] \left[\hat{\mathcal{Y}}_k^{(i)} - \hat{\mathbf{y}}_k \right]^T$
 - 23: Calculate Kalman gain (66): K_k
 - 24: Calculate posterior estimate and covariance (67)-(68): $\hat{\mathbf{x}}_{a,k}^+, P_{a,k}^+$
 - 25: Save estimation parameters for next measurement: $\left(\hat{\mathbf{x}}_{a,k-1}^+, P_{a,k-1}^+ \right) \leftarrow \left(\hat{\mathbf{x}}_{a,k}^+, P_{a,k}^+ \right)$
 - 26: Save neural network parameters for next measurement: $\left(\xi_{k-1}, \bar{\mathbf{v}}_{k-1}, (\bar{\mathbf{v}}_{k-1})^2 \right) \leftarrow \left(\xi_k, \bar{\mathbf{v}}_k, (\bar{\mathbf{v}}_k)^2 \right)$
-

For each new Mars-GRAM density profile, the random perturbations were also set to 1. Given that Mars-GRAM trajectories provide density at discrete altitudes, trajectory points were interpolated using a cubic spline to determine density values at the query altitude.

A. Monte Carlo Settings

All trajectories were initialized with the entry conditions for the Mars Science Laboratory (MSL), outlined in Table 1. This implies simulating potential MSL entry trajectories for various densities the vehicle might encounter. It is crucial to note that these trajectories are simulated in an open-loop fashion, meaning that no bank control is executed, as the focus of this work is solely on the navigation aspect. A constant bank angle ($\sigma = 0$) and a constant angle of attack ($\alpha = -17^\circ$) were assumed for all trajectories, based on the available initial entry conditions of the MSL.

Table 1 Entry conditions for the true trajectories

Parameter	Value	Standard Deviation (3σ)
r_0 (m) [17]	3.5222×10^6	3.2066×10^1
ϕ_0 (deg) [17]	-0.3919×10^1	7.8100×10^{-4}
θ_0 (deg) [17]	1.2672×10^2	3.6700×10^{-4}
v_0 (m/s) [17]	6.0833×10^3	2.6059×10^{-2}
γ_0 (deg) [17]	-1.5489×10^1	4.0000×10^{-4}
ψ_0 (deg) [17]	9.3206×10^1	2.6800×10^{-4}
B_0 (m ² /kg)	7.1000×10^{-3} [41]	4.8000×10^{-3} [10]
L_0/D_0 (n.d.)	2.4000×10^{-1} [41]	1.5178×10^{-1} [10]

The true dynamics were propagated for 350 seconds, reaching altitudes close to 11 kilometers. In addition, the dynamics were propagated with additive zero-mean Gaussian process noise and measurements were calculated using the true state with additive zero-mean Gaussian noise as well. Both the process noise and measurement noise were assumed to be white and uncorrelated with any additional sources of error. For this work, all sensors were sampled at a uniform frequency of 4 Hz. Tables 2 and 3 provide the statistics used for the process and measurement noise, respectively.

Table 2 Process noise statistics

Parameter	Standard Deviation (3σ)
r (m)	-
ϕ (deg)	-
θ (deg)	-
v (m/s)	3×10^{-1}
γ (deg)	2×10^{-3}
ψ (deg)	2×10^{-4}
B (m ² /kg)	1×10^{-5}
L/D (n.d.)	3×10^{-5}

Table 3 Measurement noise statistics

Parameter	Standard Deviation (3σ)
\tilde{a}^b (μg) [42]	3×10^2
\tilde{q} (% of reading) [23]	1
\tilde{Q}_s (% of reading)	1

B. Filtering Algorithms

To evaluate the effectiveness of the density adaptation scheme, three filters, each with different adaptive techniques, were implemented for comparative analysis. It is important to note that the focus of the comparison lies on the available adaptive methods that can be utilized online to account for the mismatch between the onboard atmospheric model and the true atmospheric density. Therefore, for a fair comparison, all adaptive methods start from the same base model, in this case the least squares exponential fit to the MARS-GRAM training trajectories. Although the exponential model is known to be a poor approximation to the true density, as it can amplify errors at higher altitudes [14], the adaptive methods are expected to account for the approximation mismatches. In addition, each solution uses identical base state spaces and measurements to enable meaningful comparison. This approach ensures a fair assessment of the strategies under consideration.

- The first filter, labeled UKF-CM (CM for covariance matching), uses the least squares exponential fit to the Mars-GRAM training trajectories for density estimation. A covariance matching technique is used to estimate the process noise covariance as defined in (20), with a batch size of 10 time steps, following the approach presented in [14]. Since the estimate of the process noise covariance may become negative definite, the diagonal elements are always reset to the absolute value of their estimates [16]. This filter is implemented using a UKF without consider analysis, since the tuning of the process noise covariance is expected to account for the uncertainty in the density.
- The second filter, denoted as UKF-AC (AC for augmentation and correction), augments the estimation state to compute a correction to a nominal profile as defined in (24) and (25). The nominal profile used is the least squares exponential fit to the Mars-GRAM training trajectories. This filter is implemented using a UKF without consider analysis as the augmentation of the state mirrors the consider approach, with the distinction that the consider parameter and its covariance are updated. The correction factor is modeled as a random walk, following similar linear adaptation approaches in previous literature [10]. This factor is initialized as the ratio between the true density and the nominal density, with an associated initial covariance of $P_{\mathcal{K},0} = 1 \times 10^{-10}$ and process noise of $Q_{\mathcal{K}} = 1 \times 10^{-7}$.

- The third filter, referred to as USKF-NN (NN for neural network), adheres to the framework detailed in Section IV, where the neural network is trained using the least squares exponential fit to the Mars-GRAM training trajectories. The network is adapted online with the maximum likelihood optimization step, and the filter uses a consider analysis to account for the uncertainty in the density estimate. The consider parameter is modeled as an ECRV centered at one with $\tau = 5$, $P_{ss} = 1 \times 10^{-3}$, and an initial covariance of $P_{c,0} = 1 \times 10^{-10}$.

The parameters used for the maximum likelihood optimization are summarized in Table 4. It is important to note that the learning rate within Adam is dynamically lowered with respect to the time step (k), as higher gradients are expected when the vehicle reaches maximum dynamic pressure. All filters use the same values for the parameters in the unscented transform, where $\alpha = 1$, $\kappa = 3 - L$, and $\beta = 2$. The initial filter state is sampled from a Gaussian distribution for each Monte Carlo simulation, with the true state as its mean and a diagonal covariance matrix following the uncertainty values as described in Table 1.

Table 4 Maximum likelihood parameters used in the USKF-NN

Parameter	Value
β_1	0.1
β_2	0.9
p_{\max}	1
$\alpha_{\mathcal{L}}$	$0.01/k$
\mathcal{T}	1

C. Monte Carlo Results

To quantify the accuracy of each filter, the time-averaged root mean squared error (RMSE) is used, defined as:

$$\text{RMSE}_i = \frac{1}{N_k} \sum_{k=1}^{N_k} \frac{1}{N_m} \sum_{j=1}^{N_m} \sqrt{(x_i(k)^{(j)} - \hat{x}_i^+(k)^{(j)})^2}, \quad (69)$$

where N_k is total number of discrete steps in each simulation, N_m is the number of Monte Carlo simulations, $x_i(k)^{(j)}$ is the i -th true state and $\hat{x}_i^+(k)^{(j)}$ is the i -th posterior estimate. Table 5 presents the time-averaged RMSE obtained for each state using the three studied filtering strategies. The USKF-NN consistently outperforms both the UKF-CM and the UKF-AC, demonstrating significantly lower estimation errors across all states. This improvement shows that the newly introduced adaptation scheme can outperform previous adaptive strategies. Notably, the UKF-AC obtains a lower estimation error than the UKF-CM for most of the estimated states, attributed to the adaptive density scheme employed in this filter.

Table 5 Time-averaged root mean squared error for the three filtering strategies

Parameter	UKF-CM	UKF-AC	USKF-NN
\tilde{r} (m)	3.0631×10^3	4.5090×10^3	6.5492×10^1
$\tilde{\phi}$ (deg)	5.7906×10^{-3}	6.7430×10^{-4}	2.3474×10^{-4}
$\tilde{\theta}$ (deg)	1.4322×10^{-1}	1.4548×10^{-2}	6.1734×10^{-4}
\tilde{v} (m/s)	0.7515×10^1	0.8712×10^1	5.5504×10^{-1}
$\tilde{\gamma}$ (deg)	3.7167×10^1	8.3381×10^{-1}	1.2342×10^{-2}
$\tilde{\psi}$ (deg)	1.1331×10^{-2}	1.5826×10^{-3}	1.1364×10^{-3}
\tilde{B} (m ² /kg)	8.1736×10^{-5}	2.0957×10^{-4}	4.3353×10^{-5}
\tilde{L}/D (n.d.)	1.8454×10^{-3}	1.7565×10^{-3}	1.7378×10^{-3}

Figure 7 shows the root mean squared error as a function of time for each estimation state and each of the three filtering strategies. From this figure, it can be seen that the USKF-NN outperforms every filter for all considered time steps, achieving the lowest estimation error across all states. In contrast, the UKF-AC and the UKF-CM show high estimation errors, particularly in states affected by the atmospheric density. The states not directly dependent on the density for dynamic propagation, such as the heading azimuth (since the bank angle is set to zero), the inverse of the ballistic coefficient, and the lift-to-drag ratio, have similar estimation errors across all the filtering strategies. There is an exception for the heading azimuth in the UKF-CM, where the estimation error quickly diverges at the start of the trajectory. However, it is important to note that the UKF-CM tends to converge faster than the UKF-AC in some of the observable states such as the velocity and inverse of the ballistic coefficient (A more comprehensive observability analysis is detailed in [29]). The comparison of these three filters highlights the significant improvement in accuracy achieved when the onboard atmospheric density is adapted using the maximum likelihood optimization in the USKF-NN scheme, especially in states directly influenced by atmospheric conditions during dynamic propagation.

The consistency improvement of the USKF-NN over the other filtering strategies can be further analyzed in Fig. 8 and Fig. 9. These figures show the error plots obtained with the USKF-NN and the UKF-AC, respectively. The error plots for the UKF-CM are not presented, as they were not considered particularly informative for the analysis, as the filter is highly inconsistent. The estimates obtained with the USKF-NN exhibit minimal bias as the estimation error remains close to zero mean for all states. In addition, the predicted covariance of this filter closely aligns with the sample variance of the errors, showing that the USKF-NN is a consistent filter for this scenario. It is important to note that the estimates for the velocity struggle to be zero mean when maximum dynamic pressure is reached (although the

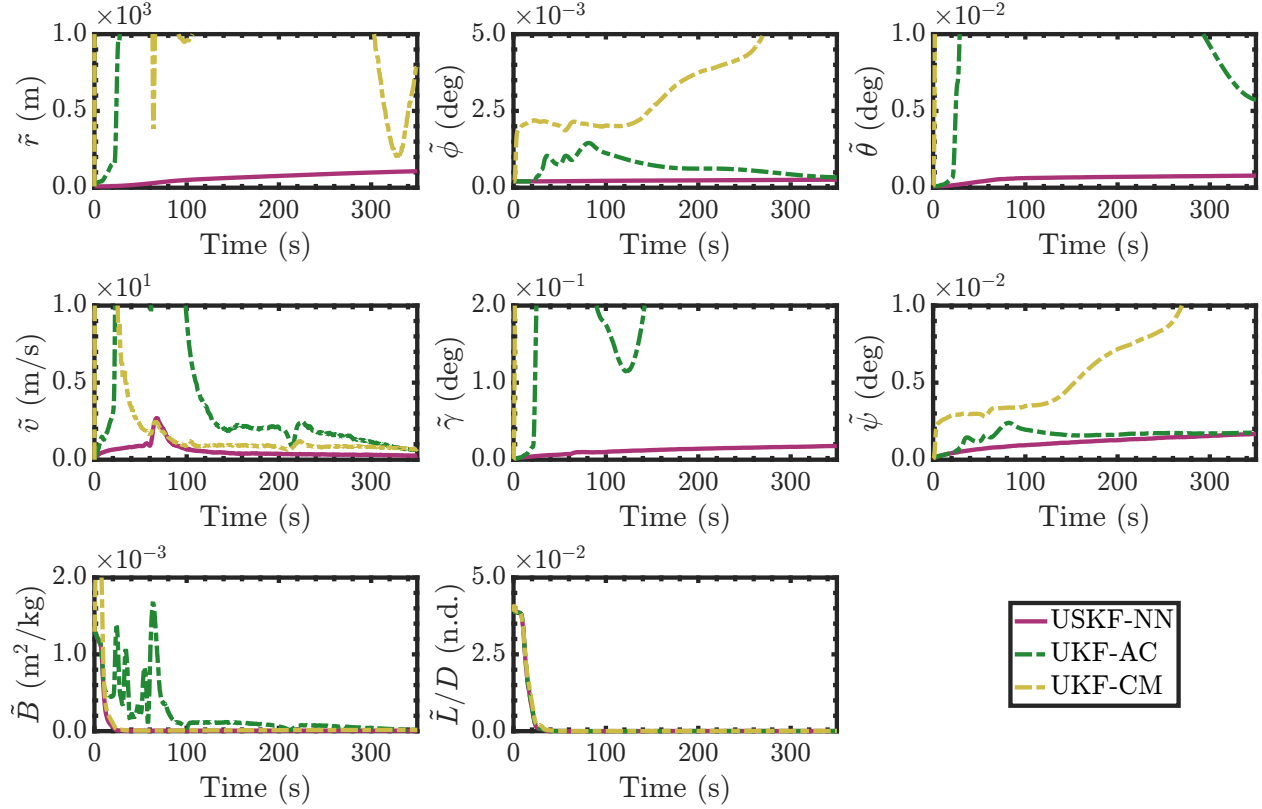


Fig. 7 Root mean squared error for each state obtained with each of the three filtering strategies.

covariance of the filter accounts for this error), but the filter quickly recovers maintaining unbiased estimates for the remainder of the flight. Conversely, the UKF-AC struggles to obtain unbiased estimates as the errors are highly biased (the estimates for the heading azimuth remain mostly unbiased since $\sigma = 0$, thus making the dynamic propagation of this state not directly affected by density). The disparity between the true atmospheric density and the onboard estimated density can rapidly introduce bias into the estimates, explaining the sub-optimal performance of the UKF-AC in this regard. This filter also struggles in capturing the uncertainty of its estimates as the predicted filter covariance does not match the variance of the estimation errors.

With the measurement setup used in this work, the estimated states are not entirely observable. Consequently, only the velocity and aerodynamic parameters appear as observable states, as seen in Fig. 8. Nevertheless, the USKF-NN achieves consistent results that can potentially be improved as more measurements become available in the descent and landing phases. The results from the UKF-AC are uninformative and it would be difficult to recuperate the estimates once more measurements are available.

Given the comparable strategies employed by the UKF-AC and the USKF-NN, both incorporating online adaptation of the atmospheric density model, it is important to quantify the error in the density estimates associated with this online adaptation. To measure the difference between the true atmospheric density and the estimate generated by each filter,

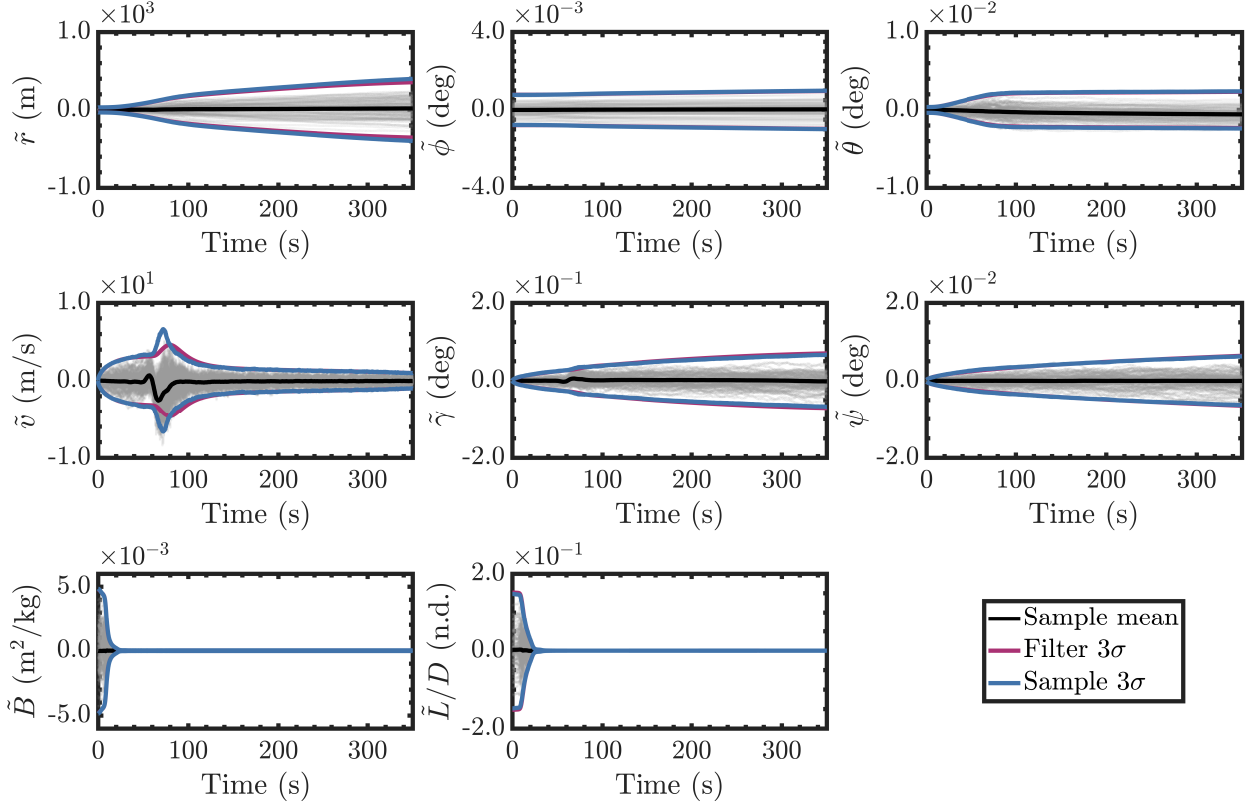


Fig. 8 Estimation error as a function of time for the results obtained with the USKF-NN. The gray lines show different Monte Carlo trajectories.

the time-averaged root mean squared percentage error (RMSPE) is used. In this work, this metric is defined as follows:

$$\text{RMSPE} = \frac{1}{N_k} \sum_{k=1}^{N_k} \frac{1}{N_m} \sum_{j=1}^{N_m} \sqrt{\left(\frac{\rho(k)^{(j)} - \hat{\rho}(k)^{(j)}}{\rho(k)^{(j)}} \right)^2} \cdot 100\%, \quad (70)$$

where N_k is total number of discrete steps in each simulation, N_m is the number of Monte Carlo simulations, $\rho(k)^{(j)}$ is the true atmospheric density and $\hat{\rho}(k)^{(j)}$ is the estimated density. Table 6 displays the time-averaged RMSPE of the density estimates for both the UKF-AC and USKF-NN. The estimates derived from the USKF-NN significantly outperform those from the UKF-AC, evidenced by the substantial difference in RMSPE. The USKF-NN excels over the UKF-AC by nearly two orders of magnitude, representing a smaller error in relation to the true density.

Table 6 Time-averaged root mean squared percentage error of the estimated densities

Filter	RMSPE (%)
UKF-AC	32.4933
USKF-NN	0.6960

The difference in density estimates between these two filters can be further analyzed in Fig. 10, which shows the

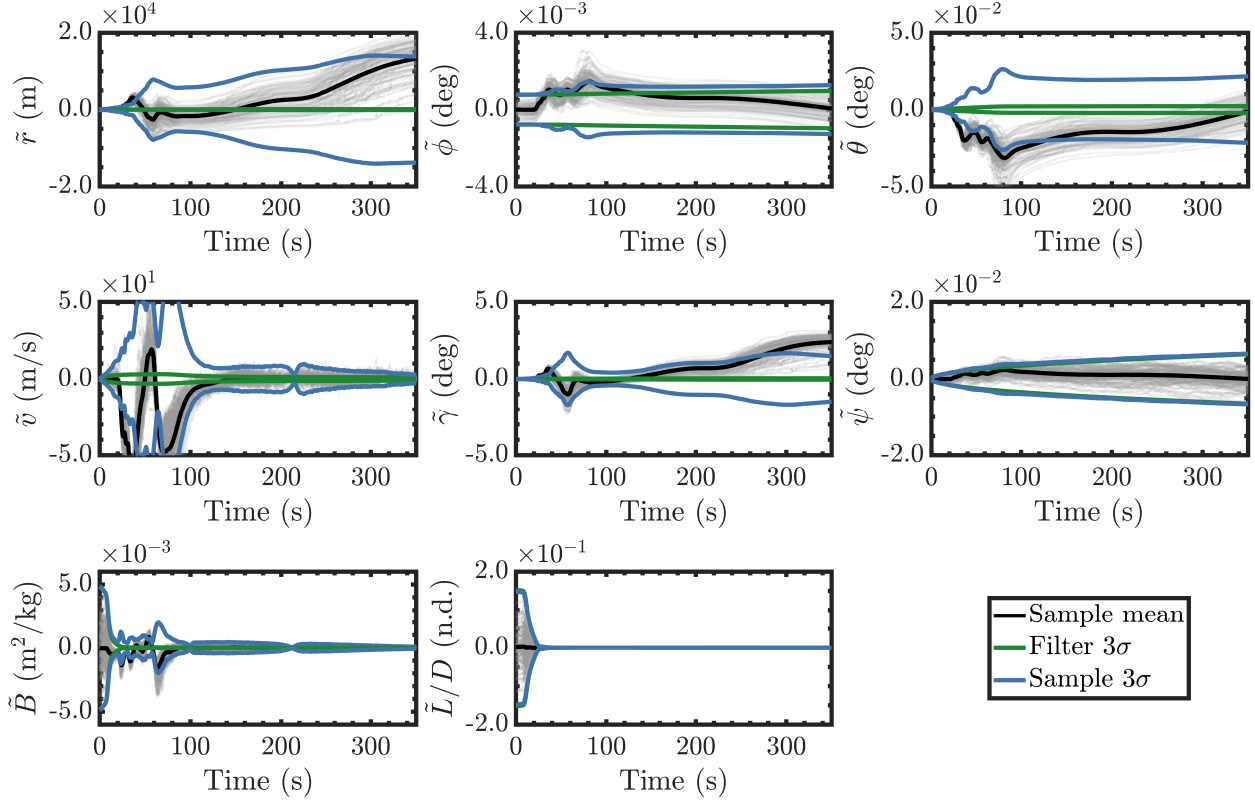


Fig. 9 Estimation error as a function of time for the results obtained with the UKF-AC. The gray lines show different Monte Carlo trajectories. Note that, for improved readability, the y-scale differs from that in Fig. 8 for variables \tilde{r} , $\tilde{\theta}$, \tilde{v} , and $\tilde{\gamma}$.

percentage error of the estimated density as a function of time for both solutions. The figure clearly shows the superior performance of the USKF-NN over the UKF-AC. The density estimates generated by the USKF-NN closely track the true density, exhibiting a very low percentage error. In contrast, the UKF-AC struggles to accurately follow the true density, resulting in relative errors that exceed the 50% mark. As previously mentioned, the USKF-NN adapts the density model in a nonlinear fashion with respect to the measurement, while the UKF-AC uses a linear update to this model. The discernible difference in error between the USKF-NN and the UKF-AC explains the improvement in estimated states, as depicted in Fig. 8, highlighting the performance gains achieved through this nonlinear adaptation technique.

VI. Limitations and Future Work

While this work has shown the potential benefits of using neural networks as an adaptive estimation method over conventional approaches, such as covariance matching and state augmentation and correction, it is essential to acknowledge certain limitations and assumptions inherent in the methodology.

In this work, the attitude of the vehicle, particularly the bank angle and angle of attack, were assumed to be

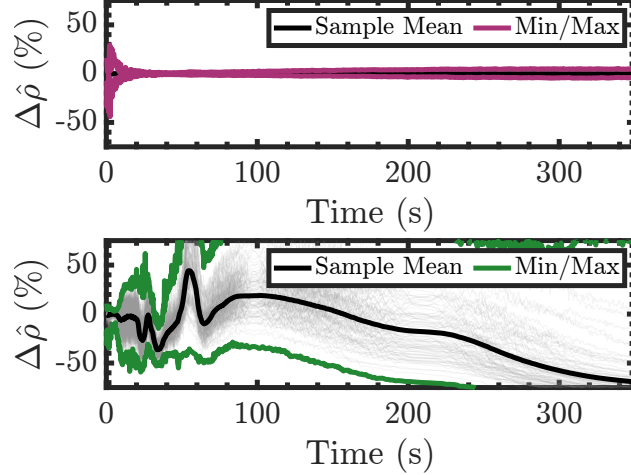


Fig. 10 Percentage error of the estimated density as a function of time for the USKF-NN (top) and the UKF-AC (bottom). The gray lines show different Monte Carlo trajectories.

constant. However, the angle of attack typically varies in entry applications [17] and bank angle control is a common guidance strategy for Martian entry missions [28]. This highlights the need to consider non-constant attitudes in future work. Nevertheless, given that the methodology described in this work is independent of control inputs, the maximum likelihood adaptation scheme is expected to remain consistent in the presence of bank angle control. Future iterations of this work will address the consideration of non-constant bank angles for the integration of the network into a closed-loop simulation.

Additionally, this work assumes that the vehicle can recover onboard the dynamic pressure and convective heating rate up to an additive white-noise term, rather than processing raw observations. More precise methods, such as Newtonian flow modeling, could be used to extend this work into a higher fidelity testing scenario. In Newtonian flow modeling, surface pressures at each port are determined using approximations based on factors like port location, angle of attack, angle of side-slip, and Mach number [43]. In this case, special attention should be given to the loss function used in (54) since the measurement noise may not adhere to a simple additive model.

In the maximum likelihood optimization step, the loss function exclusively uses the current measurement. To account for previous measurements, the optimization step is performed using the Adam optimizer, which retains memory of past gradients without incurring the time complexity increase associated with incorporating all previous measurements. However, including a window (w_o) of past and current measurements in the loss function could potentially improve filtering performance. It is worth noting that such adjustments linearly increase the time complexity of the optimization step. In the current framework, this modification would increase the time complexity of the optimization step to $O(2 \cdot h_o \cdot n_o \cdot w_o)$. Therefore, if implemented, it would be important to find an optimal window length where the accuracy gain outweighs the computational cost.

Finally, an advantage of the proposed method is that the nominal model used to train the neural network offline

can be easily changed without needing to modify the online adaptation portion. Therefore, instead of training the neural network using an exponential model, it could be trained on a higher-fidelity model that better represents the expected true density profile. Additionally, this approach is not constrained to using only one neural network. Multiple neural networks could be trained at varying altitudes using a multi-height exponential model. In this way, each neural network is adapted online when the estimated position falls within the range corresponding to that network. These modular capabilities make the presented framework a good candidate for training on higher-fidelity models before actual deployment, which could improve its robustness.

VII. Conclusions

This work introduces a new approach to onboard adaptive filtering for Martian entry using a neural network to estimate atmospheric density, and employing a consider analysis to account for the uncertainty in this parameter. The network is trained offline on a simple exponential atmospheric density model, and its parameters are dynamically adapted in real time to address any disparities between the true and estimated atmospheric densities. The adaptation of the network is formulated as a maximum likelihood problem by leveraging the measurement innovations of the filter to identify the optimal network parameters. The incorporation of a neural network in the current framework enables the use of stochastic optimizers from the machine learning domain, such as Adam, within the context of maximum likelihood adaptive estimation. These stochastic optimizers have been designed for neural networks and are known for their accuracy and efficiency.

Three filtering strategies were implemented and compared to evaluate the effectiveness of the proposed adaptation scheme. An unscented Kalman filter with covariance matching, an unscented Kalman filter with density adaptation by state augmentation, and the newly presented approach were tested through Monte Carlo analysis with different possible atmospheric profiles sampled from Mars-GRAM. The new approach outperformed the other two filters in terms of the estimation error, yielding mostly unbiased and consistent estimates. In addition, the new approach successfully adapted the onboard density model to match the atmospheric profiles provided by Mars-GRAM within an error of 0.7%.

Acknowledgment

This material is based on research sponsored by Air Force Office of Scientific Research (AFOSR) under award number: FA9550-22-1-0419.

References

- [1] Zanetti, R., "Advanced Navigation Algorithms for Precision Landing," Ph.D. thesis, The University of Texas at Austin, 2007.
- [2] Tracy, K., Falcone, G., and Manchester, Z., "Robust Entry Guidance with Atmospheric Adaptation," *AIAA SciTech Forum*,

2023. <https://doi.org/10.2514/6.2023-0301>.
- [3] Putnam, Z. R., and Braun, R. D., “Drag-Modulation Flight-Control System Options for Planetary Aerocapture,” *Journal of Spacecraft and Rockets*, Vol. 51, No. 1, 2014, pp. 139–150. <https://doi.org/10.2514/1.A32589>.
- [4] Liu, G., England, S. L., Lillis, R. J., Withers, P., Mahaffy, P. R., Rowland, D. E., Elrod, M., Benna, M., Kass, D. M., Janches, D., and Jakosky, B., “Thermospheric Expansion Associated With Dust Increase in the Lower Atmosphere on Mars Observed by MAVEN/NGIMS,” *Geophysical Research Letters*, Vol. 45, No. 7, 2018, pp. 2901–2910. <https://doi.org/10.1002/2018GL077525>.
- [5] Braun, R. D., and Manning, R. M., “Mars exploration entry, descent and landing challenges,” *2006 IEEE Aerospace Conference*, 2006. <https://doi.org/10.1109/AERO.2006.1655790>.
- [6] Lévesque, J., “Advanced Navigation and Guidance for High-Precision Planetary Landing on Mars,” Ph.D. thesis, University of Sherbrooke, 2006.
- [7] Zanetti, R., and Bishop, R., “Precision Entry Navigation Dead-Reckoning Error Analysis: Theoretical Foundations of the Discrete-Time Case,” *AAS/AIAA Astrodynamics Specialist Conference*, 2007.
- [8] Lou, T., Fu, H., Zhang, Y., and Wang, Z., “Consider unobservable uncertain parameters using radio beacon navigation during Mars entry,” *Advances in Space Research*, Vol. 55, No. 4, 2015, pp. 1038–1050. <https://doi.org/10.1016/j.asr.2014.11.016>.
- [9] Jiang, X., Li, S., and Huang, X., “Radio/FADS/IMU integrated navigation for Mars entry,” *Advances in Space Research*, Vol. 61, No. 5, 2018, pp. 1342–1358. <https://doi.org/10.1016/j.asr.2017.12.010>.
- [10] Zewge, N. S., and Bang, H., “A Distributionally Robust Fusion Framework for Autonomous Multisensor Spacecraft Navigation during Entry Phase of Mars Entry, Descent, and Landing,” *Remote Sensing*, Vol. 15, No. 1139, 2023. <https://doi.org/10.3390/rs15041139>.
- [11] Marcus, C., Setterfield, T., and Zanetti, R., “Variable Resolution Quadtree Mapping for Planetary Landing Using Planar Elements,” *AAS Guidance and Control Conference*, 2022.
- [12] Marcus, C., and Zanetti, R., “MHN-SLAM for Planetary Landing,” *AAS/AIAA Astrodynamics Specialist Conference*, 2023.
- [13] Dutta, S., Braun, R. D., Russell, R. P., Striepe, S. A., and Clark, I. G., “Comparison of Statistical Estimation Techniques for Mars Entry, Descent, and Landing Reconstruction,” *Journal of Spacecraft and Rockets*, Vol. 50, No. 6, 2013, pp. 1207–1221. <https://doi.org/10.2514/1.A32459>.
- [14] Dutta, S., Braun, R. D., and Karlgaard, C. D., “Uncertainty Quantification for Mars Entry, Descent, and Landing Reconstruction Using Adaptive Filtering,” *Journal of Spacecraft and Rockets*, Vol. 51, No. 3, 2014, pp. 967–977. <https://doi.org/10.2514/1.A32716>.
- [15] Mehra, R., “Approaches to adaptive filtering,” *IEEE Transactions on Automatic Control*, Vol. 17, No. 5, 1972, pp. 693–698. <https://doi.org/10.1109/TAC.1972.1100100>.

- [16] Myers, K., and Tapley, B., "Adaptive sequential estimation with unknown noise statistics," *IEEE Transactions on Automatic Control*, Vol. 21, No. 4, 1976, pp. 520–523. <https://doi.org/10.1109/TAC.1976.1101260>.
- [17] Dutta, S., and Braun, R. D., "Statistical Entry, Descent, and Landing Performance Reconstruction of the Mars Science Laboratory," *Journal of Spacecraft and Rockets*, Vol. 51, No. 4, 2014, pp. 1048–1061. <https://doi.org/10.2514/1.A32937>.
- [18] Wagner, J., Wilhite, A., Stanley, D., and Powell, R., "An Adaptive Real Time Atmospheric Prediction Algorithm for Entry Vehicles," *3rd AIAA Atmospheric Space Environments Conference*, 2011. <https://doi.org/10.2514/6.2011-3200>.
- [19] Amato, D., and McMahon, J. W., "Deep Learning Method for Martian Atmosphere Reconstruction," *Journal of Aerospace Information Systems*, Vol. 18, 2021, pp. 728–738. <https://doi.org/10.2514/1.I010922>.
- [20] Roelke, E., McMahon, J. W., Braun, R. D., and Hattis, P. D., "Atmospheric Density Estimation Techniques for Aerocapture," *Journal of Spacecraft and Rockets*, Vol. 60, 2023, pp. 942–956. <https://doi.org/10.2514/1.A35197>.
- [21] Zanetti, R., and D'Souza, C., "Recursive Implementations of the Consider Filter," *Journal of the Astronautical Sciences*, Vol. 60, No. 3, 2013, p. 672–685. <https://doi.org/10.1007/s40295-015-0068-7>.
- [22] Stauch, J., and Jah, M., "Unscented Schmidt–Kalman Filter Algorithm," *Journal of Guidance, Control, and Dynamics*, Vol. 38, No. 1, 2015, pp. 117–123. <https://doi.org/10.2514/1.G000467>.
- [23] Gazarik, M. J., Wright, M. J., Little, A., Cheatwood, F. M., Herath, J. A., Munk, M. M., Novak, F. J., and Martinez, E. R., "Overview of the MEDLI Project," *2008 IEEE Aerospace Conference*, 2008. <https://doi.org/10.1109/AERO.2008.4526285>.
- [24] Hwang, H., Bose, D., Wright, H., White, T. R., Schoenenberger, M., Santos, J., Karlgaard, C. D., Kuhl, C., Oishi, T., and Trombetta, D., "Mars 2020 Entry, Descent and Landing Instrumentation 2 (MEDLI2)," *46th AIAA Thermophysics Conference*, 2016. <https://doi.org/10.2514/6.2016-3536>.
- [25] White, T. R., Mahzari, M., Miller, R. A., Tang, C. Y., Karlgaard, C. D., Alpert, H., Wright, H. S., and Kuhl, C., "Mars Entry Instrumentation Flight Data and Mars 2020 Entry Environments," *AIAA SCITECH 2022 Forum*, San Diego, CA, 2022. <https://doi.org/10.2514/6.2022-0011>.
- [26] Vinh, N., Busemann, A., and Culp, R., *Hypersonic and Planetary Entry Flight Mechanics*, The University of Michigan Press, 1980.
- [27] Christian, J., Verges, A., and Braun, R., "Statistical Reconstruction of Mars Entry, Descent, and Landing Trajectories and Atmospheric Profiles," *AIAA SPACE 2007 Conference & Exposition*, 2007. <https://doi.org/10.2514/6.2007-6192>.
- [28] Lu, P., "Entry Guidance: A Unified Method," *Journal of Guidance, Control, and Dynamics*, Vol. 37, No. 3, 2014, pp. 713–728. <https://doi.org/10.2514/1.62605>.
- [29] Lévesque, J. F., and de Lafontaine, J., "Innovative Navigation Schemes for State and Parameter Estimation During Mars Entry," *Journal of Guidance Control and Dynamics*, Vol. 30, No. 1, 2007, pp. 169–184. <https://doi.org/10.2514/1.25107>.

- [30] Sutton, K., and Randolph A. Graves, J., “A General Stagnation-Point Convective-Heating Equation for Arbitrary Gas Mixtures,” Tech. rep., NASA Langley Research Center, 1971.
- [31] Benito, J., and Mease, K. D., “Reachable and Controllable Sets for Planetary Entry and Landing,” *Journal of Guidance, Control, and Dynamics*, Vol. 33, No. 3, 2010, pp. 641–654. <https://doi.org/10.2514/1.47577>.
- [32] Regan, F. J., and Anandakrishnan, S. M., *Dynamics of Atmospheric Re-Entry*, American Institute of Aeronautics and Astronautics, 1993.
- [33] Justh, H., *Mars Global Reference Atmospheric Model 2010 Version: Users Guide*, NASA, 2014.
- [34] Marschke, J., Crassidis, J., and Lam, Q., “Multiple Model Adaptive Estimation for Inertial Navigation During Mars Entry,” *AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, 2008. <https://doi.org/10.2514/6.2008-7352>.
- [35] Chu, Q. P., Mulder, J. A., and van Woerkom, P. T. L. M., “Modified recursive maximum likelihood adaptive filter for nonlinear aircraft flight-path reconstruction,” *Journal of Guidance, Control and Dynamics*, Vol. 19, No. 6, 1996, pp. 1285–1295. <https://doi.org/10.2514/3.21784>.
- [36] Kingma, D. P., and Ba, J., “Adam: A Method for Stochastic Optimization,” *3rd International Conference for Learning Representations*, 2017. <https://doi.org/10.48550/arXiv.1412.6980>.
- [37] Zanetti, R., and DeMars, K. J., “Joseph Formulation of Unscented and Quadrature Filters with Application to Consider States,” *Journal of Guidance, Control, and Dynamics*, Vol. 36, No. 6, 2013, pp. 1860–1864. <https://doi.org/10.2514/1.59935>.
- [38] Julier, S., and Uhlmann, J., “Unscented filtering and nonlinear estimation,” *Proceedings of the IEEE*, Vol. 92, No. 3, 2004, pp. 401–422. <https://doi.org/10.1109/JPROC.2003.823141>.
- [39] Wan, E., and Van Der Merwe, R., “The unscented Kalman filter for nonlinear estimation,” *IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium*, 2000, pp. 153–158. <https://doi.org/10.1109/ASSPCC.2000.882463>.
- [40] Wan, E. A., and van Der Merwe, R., *Kalman Filtering and Neural Networks*, John Wiley & Sons, Ltd, 2001.
- [41] Way, D. W., Powell, R. W., Chen, A., Steltzner, A. D., Martin, A. M. S., Burkhart, P. D., and Mendeck, G. F., “Mars Science Laboratory: Entry, Descent, and Landing System Performance,” *2007 IEEE Aerospace Conference*, 2007. <https://doi.org/10.1109/AERO.2007.352821>.
- [42] Brown, R. G., and Hwang, P. Y. C., *Introduction to Random Signals and Applied Kalman Filtering with Matlab Exercises*, Wiley, 2012.
- [43] Beck, R., Karlgaard, C., O’Keefe, S., Siemers, P., White, B., Engelund, W., and Munk, M., “Mars Entry Atmospheric Data System Modeling and Algorithm Development,” *41st AIAA Thermophysics Conference*, 2009. <https://doi.org/10.2514/6.2009-3916>.